

PCT09

## RAW SEQUENCE LISTING

DATE: 11/14/2001

PATENT APPLICATION: US/09/856,320

TIME: 14:18:51

Input Set : A:\uemura4 sequence listing.txt

Output Set: N:\CRF3\11142001\I856320.raw

ENTERED

P. 5

```

3 <110> APPLICANT: UEMURA, Hidetoshi
4      OKUI, Akira
5      KOMINAMI, Katsuya
6      YAMAGUCHI, Nozomi
7      MITSUI, Shinichi
9 <120> TITLE OF INVENTION: NOVEL SERINE PROTEASE BSSP6
11 <130> FILE REFERENCE: UEMURA=4
13 <140> CURRENT APPLICATION NUMBER: 09/856,320
14 <141> CURRENT FILING DATE: 2001-05-21
16 <150> PRIOR APPLICATION NUMBER: JP 10-347802
17 <151> PRIOR FILING DATE: 1998-11-20
19 <160> NUMBER OF SEQ ID NOS: 41
21 <170> SOFTWARE: PatentIn version 3.1
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 1301
25 <212> TYPE: DNA
26 <213> ORGANISM: Homo sapiens
28 <220> FEATURE:
29 <221> NAME/KEY: CDS
30 <222> LOCATION: (113)..(958)
31 <223> OTHER INFORMATION:
34 <220> FEATURE:
35 <221> NAME/KEY: mat_peptide
36 <222> LOCATION: (272)..()
37 <223> OTHER INFORMATION:
40 <400> SEQUENCE: 1
41 ctgccttgtct ccacacctgg tcaggggaga gaggggagga aagccaaggg aagggaccta      60
43 actgaaaaca aacaagctgg gagaagcagg aatctgcgct cgggttccgc ag atg cag      118
44                                     Met Gln
47 agg ttg agg tgg ctg cgg gac tgg aag tca tcg ggc aga ggt ctc aca      166
48 Arg Leu Arg Trp Leu Arg Asp Trp Lys Ser Ser Gly Arg Gly Leu Thr
49   -50                               -45                               -40
51 gca gcc aag gaa cct ggg gcc cgc tcc tcc ccc ctc cag gcc atg agg      214
52 Ala Ala Lys Glu Pro Gly Ala Arg Ser Ser Pro Leu Gln Ala Met Arg
53 -35                               -30                               -25                               -20
55 att ctg cag tta atc ctg ctt gct ctg gca aca ggg ctt gta ggg gga      262
56 Ile Leu Gln Leu Ile Leu Leu Ala Leu Ala Thr Gly Leu Val Gly Gly
57   -15                               -10                               -5
59 gag acc agg atc atc aag ggg ttc gag tgc aag cct cac tcc cag ccc      310
60 Glu Thr Arg Ile Ile Lys Gly Phe Glu Cys Lys Pro His Ser Gln Pro
61   -1 1                               5                               10
63 tgg cag gca gcc ctg ttc gag aag acg cgg cta ctc tgt ggg gcg acg      358
64 Trp Gln Ala Ala Leu Phe Glu Lys Thr Arg Leu Leu Cys Gly Ala Thr
65   15                               20                               25
67 ctc atc gcc ccc aga tgg ctc ctg aca gca gcc cac tgc ctc aag ccc      406
68 Leu Ile Ala Pro Arg Trp Leu Leu Thr Ala Ala His Cys Leu Lys Pro
69 30                               35                               40                               45

```

## RAW SEQUENCE LISTING

DATE: 11/14/2001

PATENT APPLICATION: US/09/856,320

TIME: 14:18:51

Input Set : A:\uemura4 sequence listing.txt

Output Set: N:\CRF3\11142001\I856320.raw

```

71 cgc tac ata gtt cac ctg ggg cag cac aac ctc cag aag gag gag ggc      454
72 Arg Tyr Ile Val His Leu Gly Gln His Asn Leu Gln Lys Glu Glu Gly
73              50              55              60
75 tgt gag cag acc cgg aca gcc act gag tcc ttc ccc cac ccc ggc ttc      502
76 Cys Glu Gln Thr Arg Thr Ala Thr Glu Ser Phe Pro His Pro Gly Phe
77              65              70              75
79 aac aac agc ctc ccc aac aaa gac cac cgc aat gac atc atg ctg gtg      550
80 Asn Asn Ser Leu Pro Asn Lys Asp His Arg Asn Asp Ile Met Leu Val
81              80              85              90
83 aag atg gca tcg cca gtc tcc atc acc tgg gct gtg cga ccc ctc acc      598
84 Lys Met Ala Ser Pro Val Ser Ile Thr Trp Ala Val Arg Pro Leu Thr
85              95              100              105
87 ctc tcc tca cgc tgt gtc act gct ggc acc agc tgc ctc att tcc ggc      646
88 Leu Ser Ser Arg Cys Val Thr Ala Gly Thr Ser Cys Leu Ile Ser Gly
89 110              115              120              125
91 tgg ggc agc acg tcc agc ccc cag tta cgc ctg cct cac acc ttg cga      694
92 Trp Gly Ser Thr Ser Ser Pro Gln Leu Arg Leu Pro His Thr Leu Arg
93              130              135              140
95 tgc gcc aac atc acc atc att gag cac cag aag tgt gag aac gcc tac      742
96 Cys Ala Asn Ile Thr Ile Ile Glu His Gln Lys Cys Glu Asn Ala Tyr
97              145              150              155
99 ccc ggc aac atc aca gac acc atg gtg tgt gcc agc gtg cag gaa ggg      790
100 Pro Gly Asn Ile Thr Asp Thr Met Val Cys Ala Ser Val Gln Glu Gly
101              160              165              170
103 ggc aag gac tcc tgc cag ggt gac tcc ggg ggc cct ctg gtc tgt aac      838
104 Gly Lys Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Asn
105              175              180              185
107 cag tct ctt caa ggc att atc tcc tgg ggc cag gat ccg tgt gcg atc      886
108 Gln Ser Leu Gln Gly Ile Ile Ser Trp Gly Gln Asp Pro Cys Ala Ile
109 190              195              200              205
111 acc cga aag cct ggt gtc tac acg aaa gtc tgc aaa tat gtg gac tgg      934
112 Thr Arg Lys Pro Gly Val Tyr Thr Lys Val Cys Lys Tyr Val Asp Trp
113              210              215              220
115 atc cag gag acg atg aag aac aat tagactggac ccaccca cagcccatca      988
116 Ile Gln Glu Thr Met Lys Asn Asn
117              225
119 ccctccatth ccacttggtg ttgtgttctt gttcactctg ttaataagaa accctaagcc      1048
121 aagaccctct acgaacattc ttgtggcctc ctggactaca ggagatgctg tcacttaata      1108
123 atcaacctgg ggttcgaaat cagtgaagacc tggattcaaa ttctgccttg aaatattgtg      1168
125 actctgggaa tgacaacacc tggtttgttc tctgttgtat cccagcccc aaagacagct      1228
127 cctggccata tatcaagggt tcaataaata ttgtctaat gaaaaaaaa aaaaaaaaaa      1288
129 aaaaaaaaaa aaa                                     1301
132 <210> SEQ ID NO: 2
133 <211> LENGTH: 282
134 <212> TYPE: PRT
135 <213> ORGANISM: Homo sapiens
137 <400> SEQUENCE: 2
139 Met Gln Arg Leu Arg Trp Leu Arg Asp Trp Lys Ser Ser Gly Arg Gly
140 53          -50          -45          -40

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/856,320

DATE: 11/14/2001

TIME: 14:18:51

Input Set : A:\uemura4 sequence listing.txt

Output Set: N:\CRF3\11142001\I856320.raw

```

143 Leu Thr Ala Ala Lys Glu Pro Gly Ala Arg Ser Ser Pro Leu Gln Ala
144      -35                      -30                      -25
147 Met Arg Ile Leu Gln Leu Ile Leu Leu Ala Leu Ala Thr Gly Leu Val
148      -20                      53 54 -15                      -10
151 Gly Gly Glu Thr Arg Ile Ile Lys Gly Phe Glu Cys Lys Pro His Ser
152      -5                      -1 1 5                      10
155 Gln Pro Trp Gln Ala Ala Leu Phe Glu Lys Thr Arg Leu Leu Cys Gly
156      15                      20                      25
159 Ala Thr Leu Ile Ala Pro Arg Trp Leu Leu Thr Ala Ala His Cys Leu
160      30                      35                      40
163 Lys Pro Arg Tyr Ile Val His Leu Gly Gln His Asn Leu Gln Lys Glu
164      45                      50                      55
167 Glu Gly Cys Glu Gln Thr Arg Thr Ala Thr Glu Ser Phe Pro His Pro
168      60                      65                      70                      75
171 Gly Phe Asn Asn Ser Leu Pro Asn Lys Asp His Arg Asn Asp Ile Met
172      80                      85                      90
175 Leu Val Lys Met Ala Ser Pro Val Ser Ile Thr Trp Ala Val Arg Pro
176      95                      100                      105
179 Leu Thr Leu Ser Ser Arg Cys Val Thr Ala Gly Thr Ser Cys Leu Ile
180      110                      115                      120
183 Ser Gly Trp Gly Ser Thr Ser Ser Pro Gln Leu Arg Leu Pro His Thr
184      125                      130                      135
187 Leu Arg Cys Ala Asn Ile Thr Ile Ile Glu His Gln Lys Cys Glu Asn
188      140                      145                      150                      155
191 Ala Tyr Pro Gly Asn Ile Thr Asp Thr Met Val Cys Ala Ser Val Gln
192      160                      165                      170
195 Glu Gly Gly Lys Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val
196      175                      180                      185
199 Cys Asn Gln Ser Leu Gln Gly Ile Ile Ser Trp Gly Gln Asp Pro Cys
200      190                      195                      200
203 Ala Ile Thr Arg Lys Pro Gly Val Tyr Thr Lys Val Cys Lys Tyr Val
204      205                      210                      215
207 Asp Trp Ile Gln Glu Thr Met Lys Asn Asn -283
208 220                      225
211 <210> SEQ ID NO: 3
212 <211> LENGTH: 1323
213 <212> TYPE: DNA
214 <213> ORGANISM: Mus sp.
216 <220> FEATURE:
217 <221> NAME/KEY: CDS
218 <222> LOCATION: (103)..(930)
219 <223> OTHER INFORMATION:
222 <220> FEATURE:
223 <221> NAME/KEY: mat_peptide
224 <222> LOCATION: (244)..()
225 <223> OTHER INFORMATION:
228 <400> SEQUENCE: 3
229 ccacatctga ctagggaagt aaggcgaagg aggcccatgg aagaaaaatc taaatgaaaa 60
231 cataagctag gagaactgag gcttcaaacc tgaagctatc ta atg agg agg ctg 114

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/856,320

DATE: 11/14/2001

TIME: 14:18:51

Input Set : A:\uemura4 sequence listing.txt

Output Set: N:\CRF3\11142001\I856320.raw

```

232                                     Met Arg Arg Leu
233                                     -45
235 aag agt gac tgg aaa tta tct aca gaa acc agg gaa cct ggc gcc cgc      162
236 Lys Ser Asp Trp Lys Leu Ser Thr Glu Thr Arg Glu Pro Gly Ala Arg
237                                     -40          -35          -30
239 cct gcc cta ctc cag gcc agg atg att ctc cga ctc att gca ctt gct      210
240 Pro Ala Leu Leu Gln Ala Arg Met Ile Leu Arg Leu Ile Ala Leu Ala
241                                     -25          -20          -15
243 ctg gta aca ggg cac gta ggg gga gag acg agg atc atc aag ggt tat      258
244 Leu Val Thr Gly His Val Gly Gly Glu Thr Arg Ile Ile Lys Gly Tyr
245                                     -10          -5          -1  1          5
247 gag tgc agg cct cac tca cag cca tgg cag gtg gcc ctc ttt cag aag      306
248 Glu Cys Arg Pro His Ser Gln Pro Trp Gln Val Ala Leu Phe Gln Lys
249                                     10          15          20
251 aca cgg ctt ctc tgt ggg gca acc ctc atc gcc ccc aaa tgg ctc ctg      354
252 Thr Arg Leu Leu Cys Gly Ala Thr Leu Ile Ala Pro Lys Trp Leu Leu
253                                     25          30          35
255 aca gca gcc cac tgc cgc aag ccc cat tac gtg atc ctc ctt gga gag      402
256 Thr Ala Ala His Cys Arg Lys Pro His Tyr Val Ile Leu Leu Gly Glu
257                                     40          45          50
259 cac aat cta gag aag aca gac ggc tgt gag cag agg cgg atg gcc act      450
260 His Asn Leu Glu Lys Thr Asp Gly Cys Glu Gln Arg Arg Met Ala Thr
261                                     55          60          65
263 gag tcc ttc ccc cac ccc gac ttc aac aac agc ctc ccc aac aaa gac      498
264 Glu Ser Phe Pro His Pro Asp Phe Asn Asn Ser Leu Pro Asn Lys Asp
265 70          75          80          85
267 cac cgg aat gac ata atg ctt gtg aag atg tgc tct ccc gtc ttc ttt      546
268 His Arg Asn Asp Ile Met Leu Val Lys Met Ser Ser Pro Val Phe Phe
269                                     90          95          100
271 acc cga gct gtg cag cca ctc acc ctg tcc cca cac tgt gtc gct gca      594
272 Thr Arg Ala Val Gln Pro Leu Thr Leu Ser Pro His Cys Val Ala Ala
273                                     105          110          115
275 ggc acc agc tgc ctc att tct gga tgg ggc acc acg tcc agc ccc cag      642
276 Gly Thr Ser Cys Leu Ile Ser Gly Trp Gly Thr Thr Ser Ser Pro Gln
277                                     120          125          130
279 ttg cgc ctg cct cat tcc ttg cga tgt gcc aat gtc tcc atc atc gaa      690
280 Leu Arg Leu Pro His Ser Leu Arg Cys Ala Asn Val Ser Ile Ile Glu
281                                     135          140          145
283 cac aag gag tgt gag aag gcc tac ccg ggc aac atc aca gac acc atg      738
284 His Lys Glu Cys Glu Lys Ala Tyr Pro Gly Asn Ile Thr Asp Thr Met
285 150          155          160          165
287 ctg tgc gcc agt gtt cgg aaa gag ggc aag gac tcc tgt cag ggt gac      786
288 Leu Cys Ala Ser Val Arg Lys Glu Gly Lys Asp Ser Cys Gln Gly Asp
289                                     170          175          180
291 tct gga ggc ccc ctg gtc tgc aac gga tct ctt caa ggc atc atc tcc      834
292 Ser Gly Gly Pro Leu Val Cys Asn Gly Ser Leu Gln Gly Ile Ile Ser
293                                     185          190          195
295 tgg ggt cag gac cca tgt gcc gtc acc aga aag cct ggt gtc tat aca      882
296 Trp Gly Gln Asp Pro Cys Ala Val Thr Arg Lys Pro Gly Val Tyr Thr

```

## RAW SEQUENCE LISTING

DATE: 11/14/2001

PATENT APPLICATION: US/09/856,320

TIME: 14:18:51

Input Set : A:\uemura4 sequence listing.txt

Output Set: N:\CRF3\11142001\I856320.raw

```

297          200          205          210
299 aaa gtc tgc aaa tac ttt aac tgg atc cac gag gtt atg agg aac aat      930
300 Lys Val Cys Lys Tyr Phe Asn Trp Ile His Glu Val Met Arg Asn Asn
301      215          220          225
303 tagaggggac ctgcttccca ccacccaacc cctccaacct cttcttaatg ctttgacttc      990
305 tcttcattct gccctaagaa gtcctcagct gggaccctgg catgtactct ctccgaccca      1050
307 ccatgagtat agtataggga tgctctaact tgatgatoga cctggggcct ggaatcaaat      1110
309 cctgacttga actaaattgt gactctggac atgatcacca ctgggtttgt ttgtttggtt      1170
311 gttttttgtt ttgttttgtt ttgttcccag ctttgaagac agtccctggc atatcccagg      1230
313 gtttcaataa atatttgta aatgataaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa      1290
315 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa                                1323
318 <210> SEQ ID NO: 4
319 <211> LENGTH: 276
320 <212> TYPE: PRT
321 <213> ORGANISM: Mus sp.
323 <400> SEQUENCE: 4
325 Met Arg Arg Leu Lys Ser Asp Trp Lys Leu Ser Thr Glu Thr Arg Glu
326      -45          -40          -35
329 Pro Gly Ala Arg Pro Ala Leu Leu Gln Ala Arg Met Ile Leu Arg Leu
330      -30          -25          -20
333 Ile Ala Leu Ala Leu Val Thr Gly His Val Gly Gly Glu Thr Arg Ile
334      -15          -10          -5          -1 1
337 Ile Lys Gly Tyr Glu Cys Arg Pro His Ser Gln Pro Trp Gln Val Ala
338      5          10          15
341 Leu Phe Gln Lys Thr Arg Leu Leu Cys Gly Ala Thr Leu Ile Ala Pro
342      20          25          30
345 Lys Trp Leu Leu Thr Ala Ala His Cys Arg Lys Pro His Tyr Val Ile
346      35          40          45
349 Leu Leu Gly Glu His Asn Leu Glu Lys Thr Asp Gly Cys Glu Gln Arg
350 50          55          60          65
353 Arg Met Ala Thr Glu Ser Phe Pro His Pro Asp Phe Asn Asn Ser Leu
354      70          75          80
357 Pro Asn Lys Asp His Arg Asn Asp Ile Met Leu Val Lys Met Ser Ser
358      85          90          95
361 Pro Val Phe Phe Thr Arg Ala Val Gln Pro Leu Thr Leu Ser Pro His
362      100          105          110
365 Cys Val Ala Ala Gly Thr Ser Cys Leu Ile Ser Gly Trp Gly Thr Thr
366      115          120          125
369 Ser Ser Pro Gln Leu Arg Leu Pro His Ser Leu Arg Cys Ala Asn Val
370 130          135          140          145
373 Ser Ile Ile Glu His Lys Glu Cys Glu Lys Ala Tyr Pro Gly Asn Ile
374      150          155          160
377 Thr Asp Thr Met Leu Cys Ala Ser Val Arg Lys Glu Gly Lys Asp Ser
378      165          170          175
381 Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Asn Gly Ser Leu Gln
382      180          185          190
385 Gly Ile Ile Ser Trp Gly Gln Asp Pro Cys Ala Val Thr Arg Lys Pro
386      195          200          205
389 Gly Val Tyr Thr Lys Val Cys Lys Tyr Phe Asn Trp Ile His Glu Val

```

Use of n and / or Xaa has been detected in the Sequence Listing. Review the Sequence Listing to ensure a corresponding explanation is present in the <220> to <223> fields of each sequence using n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/856,320

DATE: 11/14/2001

TIME: 14:18:52

Input Set : A:\uemura4 sequence listing.txt

Output Set: N:\CRF3\11142001\I856320.raw

L:975 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:36

L:1000 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:37